GFAR

GODAN
Global Open Data
for Agriculture & Nutrition

CTA

Digital and Data-Driven Agriculture: Harnessing the Power of Data for Smallholders

Preface

Knowledge is power… ICTs and the open data revolution promise a radical step change in transforming agriculture and food systems around the world, towards eliminating hunger and poverty, creating truly sustainable systems and meeting the 2030 Sustainable Development Goals. Advances in sensors, computational power and networks massively increase our ability to access, analyze and recombine big data sets, while the opening up of social media liberates the rapid exchange of information and experience. New developments, such as blockchain and big data, will shorten supply chains and quality assurance systems to bring benefits directly back to rural communities. Cellphone technology offers widespread access to data and applications. The great majority of smallholder farmers now have access to cellphones and youth are particularly well placed to capitalize on the technology.

Open data and informatics are already changing agri-food systems in all respects, from input supply to production and through to markets and consumption, overcoming barriers of cost, capacity, access, feasibility and quality. ICTs and open data will hence strongly influence the success, or otherwise of all the UN Sustainable Development Goals.

Nonetheless, opportunity for some brings real risk for others. As seen in other agricultural advances, such as plant breeding, there is no prior certainty that the poor will be able to benefit from commercially developed applications, indeed there is a significant risk that they will be further marginalized and disadvantaged as the last in line to benefit from the data revolution. Around half the world's food comes from farms of 20ha or less, but in many developing countries the data required is simply not available, the telecommunications infrastructure is weak and the vast majority of farmers cannot directly afford digital technologies for small farm plots, while women still have much less cellphone access than men. In the absence of self-sustaining markets, many start-ups fail within their first few years. Technologies and applications can open out new employment opportunities for youth, women and the marginalized, breaking down barriers of distance or access – and yet may also present a threat to some existing areas of employment, particularly for the unskilled, while the developers of digital solutions may not come from the most disadvantaged part of society. Making good use of data for the self-advancement of poor rural communities has thus become a development challenge in itself.

For all these reasons, there is a pressing need to 'democratize' access to data and technological advances for the resource poor, in forms that are both useful and usable. We must act fast and act together: no one organization or sector has the knowledge or capacity to deliver the changes required. We need to reframe and reposition data of all forms, and knowledge access and use, as core drivers of sustainable development.

GFAR, the Global Forum on Agricultural Research and Innovation, working through the GODAN collective action of hundreds of partners, embraces the opportunity that open data provides to transform agri-food systems and provide new opportunity for the world's poor family farmers. This GFAR White Paper: '**Digital and Data-driven Agriculture. Harnessing the Power of Data for Smallholders**', has been prepared by a renowned international group of experts in the field and is proudly co-published with GODAN and CTA. It sets out the opportunities for smallholders to benefit from access to relevant data in useable forms, but also the very real risks and challenges faced in enabling farmers own data – whether directly or indirectly shared - to be valued and rewarded appropriately, so that sharing data does not actually undermine their own advancement by favoring others who are better able to use it for their own purposes.

The paper contributes towards the Collective Action on Farmers Data Rights and Use being developed with GODAN, BLE, CTA, and many others and sets out in very practical terms, the tools, policies, applications and investments required for small family farms to become smart data managers and users, able to make good use of these exciting advances to create new opportunities towards each of the SDGs. Truly sustainable development must empower the poor with knowledge, realizing the benefits from data access and use and minimizing their risks, such that 'no one is left behind'. I commend this paper as a real milestone towards ensuring that vision comes true.

Mark Holderness, Exectutive Secretary, GFAR

Authors

Ajit Maru, Consultant

Dan Berne, Lagom Ag Initiative

Jeremy De Beer, University of Ottawa & Global Open Data for Agriculture and Nutrition

Peter Ballantyne, Technical Centre for Agricultural and Rural Cooperation

Valeria Pesce, Global Forum on Agricultural Research and Innovation

Stephen Kalyesubula, iLabs@Mak Project, Makerere University

Nicolene Fourie, Council for Scientific and Industrial Research, South Africa

Chris Addison, Technical Centre for Agricultural and Rural Cooperation

Anneliza Collett, Department of Agriculture, Forestry and Fisheries, South Africa

Juanita Chaves, Global Forum on Agricultural Research and Innovation

With thanks to: Mark Holderness (GFAR), Martin Parr (GODAN), André Laperrière (GODAN), Hugo Besemer (Wageningen University)

March 2018

This paper is a contribution towards the Collective Action on Farmers' Data Rights being developed through the Global Forum on Agricultural Research and Innovation (GFAR), with the Global Open Data for Agriculture and Nutrition (GODAN) initiative and the Technical Centre for Agricultural and Rural Cooperation (CTA).

Contents

Executive Summary

Information and Communication Technologies (ICT) are being used across the world to generate efficiency gains for farmers.

This has led to an information and data explosion with an associated boom in new applications, tools, actors, business models, and entire industries. Agri-food systems are being transformed.

DATA-DRIVEN AGRICULTURE – OPPORTUNITIES FOR FARMERS

Beyond the technological developments, data – on, by, of and for farmers and their products – has become a growth area, driving expectations and investments in big data, blockchain technology, precision agriculture, farmer profiling and e-extension. Investing in data-driven agriculture is expected to increase agricultural production and productivity, help adapt to or mitigate the effects of climate change, bring about more economic and efficient use of natural resources, reduce risk and improve resilience in farming, and make agri-food market chains much more efficient. Ultimately, it will contribute to worldwide food and nutrition security.

Data-driven agriculture uses big data to supplement on-farm precision agriculture – using the right farm data, at the right time and in the right formats to make better decisions. It is already being applied, big time, in many developed countries and market-oriented agri-food chains. It is being tested in many developing countries to see where data yields can be highest.

ACCESS AND SHARING – TWO MAJOR CHALLENGES

Data-driven digital agriculture offers many opportunities across the sector. It is not, however, a panacea, especially for developing-country smallholders who must overcome challenges and risks to ensure that digital investments generate dividends. For the smallholders, the two main challenges are, first, to gain access to relevant data and services provided by others and, second, to make sure that any data they share does not actually weaken their positions.

FOUR STREAMS OF DATA – BY AND FOR FARMERS

Four streams of data that farmers typically use (access or share) are identified: The first stream is 'localized' data generated and collated on the farm for use only on the farm. The second stream is 'imported' data generated and collated off the farm, for use on the farm. The third stream is 'exported' data generated and collated on the farm for use off the farm. The fourth stream is 'ancillary' data generated and collated (on and) off the farm, mainly for use off the farm.

The opportunities, challenges and risks for farmers are different in each stream. Localized and imported data, when clearly focused to on-farm situations underpin and drive good decision making on farms. From a farmer perspective, data shared (exported) by farmers locally and globally should provide 'win-win' results in terms of more relevant services and products (from forecasts to logistics to advice) they a farmer can access. 'Localized' and 'ancillary' data streams don't present special access challenges to farmers as they are either completely inside the farm or completely outside. 'Imported' data presents all the challenges of availability, accessibility and usability, with data and services offered needing to be adjusted to what farmers need and can handle. 'Exported' data streams present all the risks and benefits around sharing, with added safeguards necessary to avoid exploitation of what the farmers share.

DATA INTERMEDIARIES AND INTEGRATION

The providers, enablers and handlers of data-driven services for and with farmers are critical actors in agri-food data systems. They find and transform raw data into actionable information and decision-making tools. They need to know the real needs of the farmers on the one hand and to find and understand the necessary data on the other. They must understand the standards, formats and licenses as well as the data collection practices including measurements and biases and make it useful. They need to win the trust of data providers and users, safeguarding ownership and striving for equitable access. Data standardization is one of the biggest challenges these data intermediaries face and different pathways to developing standards can be followed.

CHALLENGES FOR SMALLHOLDERS

Smallholders are tough to reach and thus many initial data service offerings have been designed for larger and more commercial operations. Following a 'bottom of the pyramid' logic, smallholders have much to offer any service providers and intermediaries like farmer organizations who are able to design

and deliver data-driven services at large scale. Smallholders also have much to gain from data – small improvements in their operations are likely to provide larger gains at household level, proportionally, and, if the improvements are widely adopted, the whole agricultural sector in many countries that depend on smallholder agri-food systems can be transformed.

However, for smallholders to benefit from data-driven agriculture, tools and applications need to be designed for their specific situations and capacities; they – and the organizations that support them – need to grow their capacities to become smart data users and managers; measures are needed to ensure that farmer-generated data is not exploited or misused; and smallholders, usually the least powerful parts of a value chain, must grasp every opportunity to be included in the collective data flows within agri-food systems.

FACILITATING DATA USE BY SMALLHOLDERS – DRIVERS FOR CHANGE

Data-driven agriculture offers opportunities and poses threats to smallholder farmers. Making data-driven agriculture smallholder-friendly should be guided by two sets of drivers.

Important agri-food system drivers that determine the effectiveness of data-driven improvements that need to be tackled include: developing appropriate policies and related institutions and structures; devising incentives that deliver benefits to smallholders; developing capacities of farmers and small and medium entrepreneurs and institutional capacities to manage support systems for data and information sharing and exchange; extending the availability and affordability of hardware, software and data; developing needs-based data and software; and providing necessary infrastructure and connectivity within the reach of smallholders.

Important data system drivers that need to be factored into investments in this area include: developing 'apps' for farmers that enable localized and specific solutions; using open agricultural data and standards that facilitate transparent, equitable and wide use and re-use of data; using ICTs and data to create jobs and make agriculture attractive to young people; investing in new data handling technical developments like big data, blockchain and Internet of Things that will make data value chains more powerful and transparent; strengthening institutions that enable equitable governance of data, locally to globally; and promoting joint actions on data by smallholders and their representatives through farmer organizations, cooperatives, associations, enterprises, etc.

DEVELOPING SMALLHOLDER-FRIENDLY DATA ECOSYSTEMS

The final section of the paper presents three priority actions to help develop a data ecosystem to support smallholders.

- First, farmer data and services based around data should be aggregated through joint action that empowers and gives voice to farmers;

- Second, trust centers, platforms and mechanisms that enable open data sharing should be established at different levels;

- Third, international agreements to facilitate data access, ownership and flows should be developed.

While many different actors need to be involved in these action areas, there are especially clear and present opportunities for groups (associations, federations, cooperatives, social enterprises, etc.) that represent and aggregate smallholders to step up their digital investments and capabilities so their members can grasp the opportunities offered by data-driven agriculture.

Introduction

A more data-driven, information-rich and knowledge-intensive agriculture industry is rapidly evolving across the world.

This paper's starting point is that new approaches to agriculture must balance the need of producers to participate in globally competitive agricultural commodity markets while also taking account of increasing pressures on land for non-agricultural uses and working with agricultural resources that are not suitable for optimal production in many parts of the world.

Data-driven agriculture informs decisions and guides choices to help find the right balances among these different goals. It also helps people adapt to or mitigate the effects of climate change, bring about more economic and efficient use of natural resources, reduce risk, and improve resilience in farming, agriculture and agri-food market chains. Ultimately, it contributes to worldwide food and nutrition security, ensuring that "all people, at all times, have physical, social and economic access to sufficient, safe and nutritious food that meets their dietary needs and food preferences for an active and healthy life."[1]

Data-driven agriculture is enabled by multiple, diverse communication technologies, such as sensors, mobile devices, blockchain, big data analytics, the Internet of Things, cloud-based computing, web services etc. While opening up many new possibilities, the accessibility, management and use of data by and for farmers faces two main challenges: First, farmers are challenged to gain access to relevant data and services provided by other actors (so called 'imported data'). Second, sharing their own data ('exported data') opens farmers up to potential risks.

Overcoming these challenges requires measures to support more equitable sharing and exchange of data and information and steps to enable its effective use. Capacity development can help farmers, farm workers and others to more effectively use data and information. Timely and easier access to financial and market related information and data on farm inputs and logistics is important as well as clear policies on data ownership and awareness raising on data rights.

This paper argues that these challenges are particularly critical for smallholder farmers in developing countries because they lack many of the necessary capacities as well as the necessary enabling and supporting environmental factors. It also argues that these challenges can be overcome by deliberately making data-driven agriculture more smallholder-friendly.

1 Committee on World Food Security. Global Strategic Framework for Food Security & Nutrition. http://www.fao.org/cfs/cfs-home/products/onlinegsf/en/

Data Driven Agriculture

Data-driven agriculture is "the thoughtful use of big data to supplement on-farm precision agriculture. It means having the right farm data, at the right time, to make better decisions."[2]

Precision agriculture is more specifically the use of data that has been generated on the farm. It is distinct from data on or from wider value chains that can be useful to a producer but is collected, compiled or distributed by others (market information for instance). This distinction between data generated on and off farms, by and for farmers, and ways the two forms of data are combined, is a key theme in this paper.

These concepts are closely related to notions of digital farming and agriculture 4.0.[3] An extension of this concept is to use data and information flows within and across agri-food chains to improve efficiency of the chain. This is normally seen as ways to maximize profitability, however it can go beyond this and holistically optimize productivity of an entire chain or a chain component, for example by integrating for sustainability in the use and recycling of natural resources, improving product quality, enabling tracking and tracing of farm inputs and products, or reducing drudgery involved in farming.

Agricultural systems and data flows

Farming and agri-food systems are complex, dynamic systems with hundreds of component sub-systems, each with multiple actors and stakeholders, each with different roles and expectations. These systems spread horizontally –input suppliers, producers, processors and traders to consumers – and they vertically integrate with processors transforming commodities to products for different market chains. Underlying these systems is the essential data and information on the natural resources base that underlies all forms of agricultural production.

These agri-food systems are bound and act together through intertwined flows and interactions of finance, commodities and information starting from the natural agricultural resource capability and potential to farm inputs to consumption or wastage within and across the chains. Optimizing the advantages and constraints in such systems is critical to make them work, calling for sophisticated tracking and tracing of resources and flows.

Data and information flows are intertwined with these agri-food system financial and commodity flows. While commodities flow from input to output through various value adding stages in the agri-food system, finance flows from output to inputs. Data and information flows are as in a network. Each value addition stage has different types and uses of data and information. Various actors and stakeholders use, generate, process and manage data and information as per their needs and capacities. Imperfect flows, of data and information along with those of finance and commodities, can cause turbulence and failure in agri-food related markets. Transparency of these flows with more inclusive and open access to these data can help mitigate the risks of such turbulence and failure.

Data is thus both a critical input as well as a valuable product in the modern agri-food system. Open data enables and empowers. Open data, if applied flatly to address both farmer access and sharing challenges can also, if collected and applied without clearly defined principles rules, and ethics, make farmers even more vulnerable to the asymmetries of financial, commodity and information flows in agri-food chains, undermining farmers' livelihoods.

Four Streams of Data for Farming

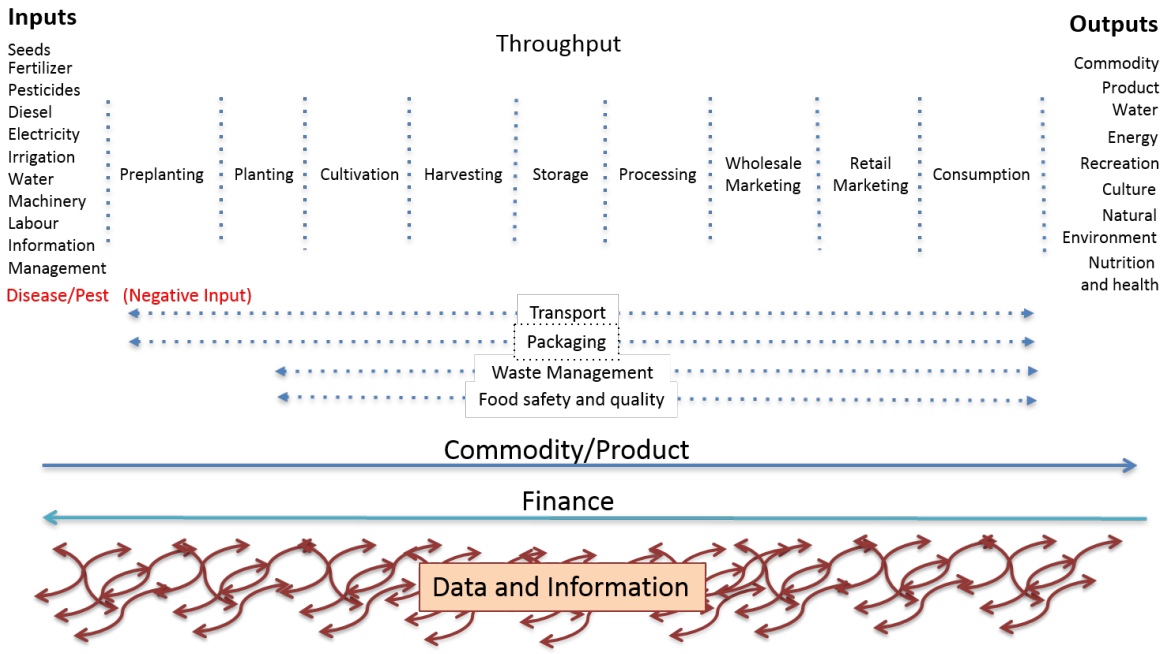
Farmers typically use (access or share) three streams of data and information in their farming. A fourth stream of off-farm data also exists but has little use for farmers.

The first stream is data generated and collated on the farm for use only on the farm. This is called 'localized' data. This includes soil data (soil form; soil depth; nutrient composition), seed and fertilizer use, date of sowing, production practices, water use, etc. that farmers have about their immediate location. This data can be generated and managed by the farmer or by an agent acting for the farmer. It would normally be 'owned' by the farmer. Some types of on-farm data and

2 Hayden, B. 2015. What does data-driven farming mean? https://blog.heatspring.com/what-does-data-driven-farming-mean

3 CEMA aisbl - European Agricultural Machinery. 2017. Digital Farming: what does it really mean? http://www.cema-agri.org/page/digital-farming-what-does-it-really-mean

Figure 1
Data and Information Flows in Crop Agri-Food Systems



information cross over into traditional and indigenous knowledge systems with their own dynamics.

The second stream is data generated and collated off the farm, for use on the farm. Examples are climatic data and market prices that have been interpreted and customized for on-farm use. This is called 'imported' data. This data is usually owned, managed and controlled by a third party and made available, directly or through intermediaries, to farmers (and their representatives). Farmers do not 'own' this data, unless they have purchased it; they have permission to exploit it. These services are often private or semi-private and they seek effective business models which often include explicit sales of data, usually packaged for clusters of customer groups.

The third stream is data generated and collated on the farm for use off the farm. This is usually processed, aggregated or combined with other data and information generated elsewhere and is used by various actors and stakeholders, such as governments or private companies. Governments may use this data to establish ownership, revenue and subsidies or to target services – that can all directly benefit farmers. Other users include market intermediaries, farm input and service providers including banks, insurance agencies, farm advisory services, scientists, other farmers and their associations etc. They each have their own need for data 'from' farms. This is called 'exported' data. While collected from farmers (or their farms using sophisticated tools like drones or remote sensing), this data and the products and services it generates are usually owned by a third party. It can be made more widely accessible by making it 'open'. There are growing concerns to safeguard this data for farmers to ensure it is not exploited at cost to the farmer. Some data collectors, such as scientists, have adopted ethical guidelines to make sure the data they collect does not exploit the farmers who provide it.[4] Some governments similarly regulate some data collection to counter perceived data piracy (and are using the Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization, for example, to safeguard the rights of their citizens). Similarly, some private companies have procedures that copy data they collect

[4]See: Chaves Posada, J. 2013. Achieving Farmers Rights in Practice: GFAR Discussion Document. Rome: Global Forum on Agricultural Research. http://www.gfar.net/sites/default/files/cgiar_farmers_rights_report_final_aug_13.pdf; and CGIAR. 2012. CGIAR Principles on the Management of Intellectual Assets. Montpellier: CGIAR http://hdl.handle.net/10947/4486

(through smart equipment and sensors for instance) to data repositories that the farmers can own and access.

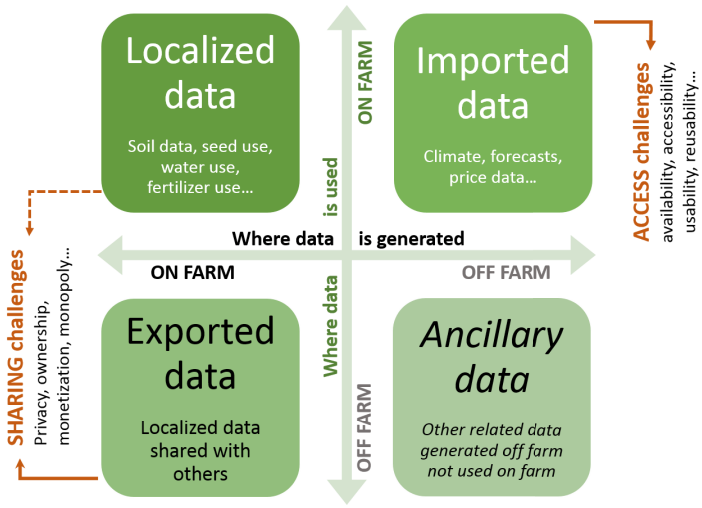
The fourth stream is data generated and collated (on and) off the farm, mainly for use off the farm. A large proportion of 'agricultural' data such as government statistical and research data are generated using various data sources, in another value chain component perhaps, and may have little direct on-farm application – though they can, through policy changes perhaps, have indirect on-farm influence. This is called 'ancillary' data. It typically has many owners and platforms and associated products. Some parts of this may find their way to farmers, directly through 'importing' services or indirectly via other actions and changes that may influence what happens on farms.

This distinction, depicted in Figure 2, between sets of data according to whether they're generated on or outside the farm and according to the potential use by different actors will be used again in the document, as the different streams present different challenges.

Looking at the two access and sharing challenges described earlier, completely localized and ancillary data streams don't present special challenges to farmers. Since they are either completely inside the farm or the community or completely outside, the issue of 'access' is irrelevant. Imported data streams clearly present all the challenges of availability, accessibility and usability. Exported data streams present all the risks and benefits related to the sharing of one's own data in public.

It is important to note though that localized data streams, when digital, are always likely to be exported. Depending on how farmers collect and manage and share their data, it can be exported (willingly or not) in various ways.

Figure 2
Streams of Farming Data



Uses for Data in Farming

Data is used for different purposes in agri-food systems. For farmers, these purposes include:

• **Planning**
What to produce, when to produce, where to produce? For whom to produce? What operations to do when and where on the farm? This needs both imported data from outside the farm (e.g. weather data, market data, crop and animal growth models) as well as localized data from the farm (e.g. soil).

• **Monitoring and assessment**
How is the product growing? What is the status of the natural agricultural resources? This requires mainly localized data from the farm (e.g. monitoring data from sensors). This is also one of the areas where data from the farm has an 'export' value for other actors (e.g. for monitoring the use of land and natural resources or for national maps of land capability).

• **Event management and intervention**
Which action should be taken and when? This requires mainly localized data from the farm (e.g. soil data from sensors) but can benefit from external data like weather forecasts, growth models or market conditions.

• **Autonomous action through ICTs**
For example, switching on water pumps to irrigate fields when soil humidity falls below a target amount, opening or closing windows in glasshouses, or auto feeding animals at different times of the day. This requires similar data as event management above.

5 De Beer, J. and Wunsch-Vincent, S. 2016. Appropriation and intellectual property in the informal economy. IN: Wunch-Vincent, S. and Kraemer-Mbula, E. (eds). 2016. The informal economy in developing nations: Hidden engine of innovation? Cambridge: Cambridge University Press. https://ssrn.com/abstract=2665172

Data, information and knowledge have always been central to farming and agriculture.

The Nilometer at Aswan and other places on the course of the Nile was a data-gathering tool that measured water clarity and levels in ancient Egypt. The priests and the Pharaoh used it to measure the Nile river floods, predicting crop performance. It also helped estimate the taxes that farmers paid to the Pharaoh and the priests.

Similarly, Stonehenge and the Mayan Pyramids and Calendars were also data tools that helped predict farming events. Again, the priests controlled the data and the analysis. Almost all ancient civilizations that grew around farming had data and information management systems embedded in their social and cultural practices.

• **Optimization**
What will be the economic, environmental or social return/effect on the investment/action? This depends completely on imported data, like market data, consumption statistics, land and water use, potential payment schemes for environmental services.

• **Forecasting**
How much will be the crop or animal yield? How much profit? This requires a combination of localized and imported data as well as prediction models: the more data the farmer (or the farmer's adviser) has, the more accurate the forecast is.

• **Tracking and tracing**
Where is the product, item, resource or material? What is its source and where will it go next? This is a crucial area where a lot of external tracking data benefits the farmer; the data shared from the farm can also become essential for the overall tracking data flows (e.g. farm identification, farming practices, agricultural input used).

• **Negotiating and market access**
Where are the consumers? What do they want? Who else is selling the same product? Which market is surplus and which glut? Which service providers can I best work with? Localized data on farm operations and products, and input needs, can be combined with imported data by farmers (or their representing organizations) to negotiate better prices, discounts, and the like. Localized 'metadata' on farmers and their farms can be powerfully aggregated by farmer organizations using joint actions to negotiate better deals for their members.

Imported data challenges for farmers

Farmers face different challenges for different streams of data.

The fundamental differences among the streams are whether data is generated on or off farms and whether it is used on farms or more widely. Localized data (on-farm for on-farm use) present the least challenges for the farmer, apart from any costs of equipment, software or methodology to collect and interpret it.

Imported data (off-farm for on-farm use) present all the challenges related to availability, accessibility and usability. More specific challenges like timeliness are especially relevant for certain types of data. The data collected, managed and used in farming can be static, such as land ownership and farm field boundaries, or dynamic, which changes over time. Some dynamic data can have very short (daily weather data) or very long lives (soil nutrient values).

Some 'imported' data challenges are inherent to the data itself and the way it's provided, while others depend on users' capacities and enabling environment.

Challenges inherent in the data include:

• **Access challenges**
a. Availability and accessibility: availability is the existence and findability of the data, while accessibility measures the extent that data can be opened (protocols, licenses, permissions) and read (formats, encoding).
b. Interoperability is the extent that data can be parsed and understood across platforms and systems (formats, standards, semantics).
c. Reusability is the extent that data can be combined with other data and reused (licensing, metadata, provenance).

• **Usefulness** depends on the purposes for which data is used. The same data can be useful to some users and not to others. The key attributes influencing data usefulness are:

a. Accuracy
b. Scale
c. Timeliness
d. Trustworthiness
e. Relevance

Challenges related to users' capacities and the enabling environment include:

• **Affordability** is the costs of access and use of data and services. Costs of access are closely related to the ability to pay of farmers, especially the poorer ones.

• **Applicability** concerns the relevance of data and services to the specific needs and capabilities of farmers. It depends partly on a farmer's capacities to apply the data but also on the extent to which data and service providers aim at making data applicable to farmers' needs. For data to support farmers, data and service providers must design or transform data so it responds to farmers' needs and situations.

• **Appropriation** applies to the capacities of farmers to 'appropriate' or take ownership of data and information in a collective manner. Appropriation may be achieved via ownership (property rights) but there are other ways in which economic, social, cultural benefits can be appropriated.[5]

• **Effective** use concerns the abilities of farmers to find, understand, interpret and use data and information effectively. A significant issue is to identify exactly which available data is useful and can be used. Some data is useful but cannot be used because it is in inappropriate formats or the farmer is unable to apply or make sense of it. Sometimes, data out of context is misinterpreted or misused because critical details about why it was created are not communicated.

For farmers, data needs to be transformed through different processes into information, and this information used with experience as knowledge. At every stage, an important core process is learning. This is not linear and, at each stage, farmers use, share and exchange data, information, and knowledge from several sources, each with their unique characteristics and applied or adapted based on the farmer's own

Recent CTA work with farmer organizations mapped three main result areas where data-driven services for, from and by farmer organizations are taking place:

Data-driven services and products that enhance on-farm PRODUCTION include:

• Accessing diagnostics and advice on areas like: agro-climatic forecasts; agronomic advisory recommendations; soil-water; pests and diseases

• Accessing early warning on threats though alerts services

Data-driven services and products that enhance access of farm products to TRADE and MARKETS include:

• Accessing markets and customers: product certification; product tracking and traceability; market information – supply, demand, competition, prices;

• Sourcing knowledge, inputs, advice: knowing value chain actors, networks, expertise and the resources, products and services they provide.

Data-driven services and products that enhance access of farmers to FINANCE include:

• Accessing financial services: banking; insurance; credit; money transfer and microfinance

Source: Addison, C. and Msengezi, C. 2018. Farmer organizations and precision agriculture data services. ICT Update (86): 16.

5 De Beer, J. and Wunsch-Vincent, S. 2016. Appropriation and intellectual property in the informal economy. IN: Wunch-Vincent, S. and Kraemer-Mbula, E. (eds). 2016. The informal economy in developing nations: Hidden engine of innovation? Cambridge: Cambridge University Press. https://ssrn.com/abstract=2665172

knowledge base. These learning and adaptation capacities are critical for farmers to benefit from data-driven agriculture opportunities.

Exported data opportunities and challenges for farmers

Exported data that has been generated on farms for use outside farms (perhaps combined with other external data) presents different challenges as well as potential benefits.

To participate in globally competitive markets, farmers need to manage vast amounts of data and information even for one crop on a single plot, field or farm. This data and information has different value, benefits and risks when it is used on an individual farm, field or plot and when aggregated across and within farms. Data generated by farmers and shared either voluntarily, through mandate, coercion or unknowingly with others, including neighboring farmers, governments, input providers, market intermediaries and consumers can have both beneficial and detrimental effects for farmers.

When shared with other farmers, crop-related information can help deliver joint actions against threats and risks like diseases. It can also be used to achieve more informed participation in markets. But, the same data can be abused by third parties to control access to markets, specific products, or to manipulate prices.

Data shared by farmers, such as planting dates for crops, when used together with data from external sources such as rainfall data can be used by irrigation managers to release water. There are many other examples like this. In India, for example, rainfall patterns are carefully monitored by gold traders, white goods and tractor manufacturers as a good rainy season at the right stages of the cropping cycle can mean higher incomes for farmers which they may spend on consumer and industrial goods.

BENEFITS TO FARMERS FROM DATA EXPORTING AND SHARING

The benefits from data sharing (exporting) for farmers locally and globally mainly result from having access to more aggregated and realistic data to inform services (from forecasts to logistics to value chain tracking) that will benefit farmers. Some examples are:

- **Collective services**
Sharing data among farmers can help build joint services and resources, such as shared machinery, tools and labor; better services through route mapping and scheduling; improved bargaining in market participation both for inputs and produce and for logistics such as transport and storage.

- **Better forecasts**
Sharing data on land and water use, crop growth, production, pests etc. can contribute to better monitoring of natural resources, better market related forecasts, forecasts for disease and pests outbreaks and subsequent shared management of outbreaks through joint action.

- **More realistic solutions from research**
Sharing data with researchers can help them generate more realistic solutions for a wide range of problems.

- **More transparency, especially in value chains** Sharing farm data like farm identification, farming practices, agricultural inputs and pests is essential to improve supply chain data tracking. Sharing data can also increase transparency in legal matters: for instance, sharing data about land tenure can mitigate conflicts related to land ownership and land grabbing.

RISKS TO FARMERS FROM DATA EXPORTING AND SHARING

The risks or abuses farmers face when they share (export) their data include[6]:

- **Misuse**
Data provided by farmers without proper safeguards can lead to issues of privacy, security, safety, or liability. Data can also be misused by third parties for their own gain. Such risks need to especially be guarded against through farm profiling where data is gathered about a farm, its inhabitants and their activities. As all sort of precision and digital tools get widely used, this may happen without the full knowledge and consent of farmers.

- **Monetization**
There is a market for data provided by farmers; it can be sold and bought. This needs to be made clear to farmers

[6]See: Ferris, L. and Rahman, Z. 2016. Responsible data in agriculture. Wallingford: GODAN. https://f1000research.com/documents/6-1306. See also the summary of an e-discussion in 2014: Chaves Posada, J. 2014. Rights of farmers for data, information and knowledge. Rome: Global Forum on Agricultural Research. http://www.gfar.net/sites/default/files/rights_of_farmers_for_data_information_and_knowledge.pdf

during data collection so they are aware what might happen with their data and who may benefit financially from it. Reusing data for commercial purposes - in many instances by third parties and perhaps not returning benefits to the original data generators - raises all sorts of ethical and moral issues.

• **Unfair competition**
When farmers provide data to another actor, this can give that actor more knowledge and a privileged position to sell tailored services – that hopefully benefit or at least don't damage the farmer providing the data. This data can also be exploited in different ways to deny services or gain unfair advantages.

Data Intermediaries and Data Integration

Farmers rarely deal with raw data, so their challenges are usually more with the information services and decision-making tools than with the data itself. The providers and handlers of data-driven services tend to face most issues with the data itself. This is true both when reusing external data on the farm and when sharing data with other actors. In data streams that import external data to the farmer, data intermediaries have critical roles in finding and transforming raw data into actionable information and decision-making tools. They need to know the real needs of the farmers on the one hand and to find and understand the necessary data on the other. They must understand the standards, formats and licenses as well as the data collection practices including measurements and biases and make it useful.

In data streams that export farm data to other segments of the value chain, intermediaries need to be aware of data ownership issues, anonymization techniques and the standards and formats to use to make data reusable by others. It is important that data is collected and exposed in standardized ways that can be re-used along the value chain.

Other important intermediaries are extension agents and commercial advisers who need to be able to interpret the data (more often at the stage of information) and to go even beyond and transform it into knowledge.

In a world of distributed data all sorts of novel services are emerging to transform and customize data and produce tailored services for different users. These involve many intermediaries at different stages (for example from the developer that integrates weather data in a crop planning tool to a crowdsourcing weather service getting data from many weather stations to a national MET Office getting global data and repurposing it) all dealing with different standards and practices, all in turn repackaging the data in still other standards and formats and so on. These data challenges of intermediaries and service providers reflect, of course, on the quality of the services they provide and therefore become indirect data challenges for farmers – who need to choose among a diverse array of services and service providers.

DATA INTEGRATION AND DATA STANDARDS

The issue of standardization is one of the biggest challenges for data intermediaries. As nodes in the data value chain, intermediaries need to understand the way data in a certain database or system are exposed and in turn be able to expose data in standardized ways for others to use.

Already at the level of precision agriculture data needs to be transferred between machines. From crop planning to farm management to inventory and supply chain tracking, agriculture is ever more dependent on the automated capturing, transfer, and management of data. Integration issues are a key barrier, as more and more farm activities rely on the capture and analysis of multiple sets of data from multiple sources.

For suppliers, using open standards removes consumer barriers around multiple systems and increases the likelihood that consumers will purchase their products and services. They save time and money on software updates, they tap into and enable the 'power of the crowd', and they can create new market opportunities.

Although the drive towards open standards is not yet a smooth one as there are still dominant proprietary standards and there is already a plethora of standards that need to be sorted out and made more usable, work on open standards is taking place in almost all segments of the data value chain and this currently offers major opportunities to achieve data integration and compatibility.

Data and the Smallholder Farmer

Categories of Farmers

Farmers are not a homogenous group. They manage large or small areas of land, apply different production methods, and grow crops or raise animals for diverse purposes. They may be organized as commercial entities, be run by individuals or join forces through cooperatives and collectives. Headed by men and women, farm families may be literate or illiterate, have access to education and other income chances, are close or far from markets and digital connections. Each thus has different issues and opportunities around data and information.

This diversity is powerful, enhancing resilience and contributing to richer, more sustainable agri-food systems. The diversity is also a major challenge to data-driven service providers whose business models need to balance delivery of precise and personalized services with the need to reach many customers at low costs.

A recent review[7] of farm size and productivity concludes that most of the world's farms are small, even in rich countries. The greatest difference across countries is in average labor productivity. The richest ten percent of countries are fifty times more productive, on average, than the poorest ten percent of countries. Increasing labor productivity in smallholder farming in developing countries looks like a promising area for data to be applied.

Challenges for Smallholders

To grasp the opportunities of data-driven agriculture, smallholder farmers in developing countries need to overcome some specific challenges.

Farmers with poor rural infrastructure such as internet, cell phone access and road connectivity, markets and capacities to use digital technologies are at the lowest level in ability to manage and effectively use data, new information and knowledge in their farming and related market participation.

As such, they are significantly more affected by unequal or insufficient information, unpredictable environmental changes in rainfall, soil erosion, scarcity of human labor with increased cost, yield loss due to pests and insects, increase in cultivation costs, as well as poor and perhaps exploitative supply and market chain management.

Larger, more commercial farmers with access to better infrastructure and supporting farmer organizations are, generally, better placed on these issues. Generally, the ability to manage and effectively use data and information reflects the power of farmers in agri-food systems and their participation in market chains.

In the United States, farms are categorized by revenue, not land size. The US Department of Agriculture categorizes farms primarily on the basis of Gross Cash Farm Income which includes the farm's sales, receipts of government payments, and other farm-related income.

Adapting to changes in the national agri-food system, supersized family farms are growing to stay profitable, squeezing the revenue and profits of smaller operations. Just 4% of US farms produce 67% of the agricultural output. And just 12% of farms control 75% of US cropland.

Larger size seems to be linked to higher adoption of precision agriculture. A number of reasons for this include:

Cost and affordability for larger operations

Greater efficiencies and return on investment over a larger site or multiple sites

Availability of expert or trained personnel, some of whom can use their expertise over multiple sites.

Higher level of support from vendors,

The implication for smallholders is that they need to find ways to obtain the same type of adoption advantages. Perhaps through support from governments, NGOs, agents and others or by pooling efforts through associations.

[7]ISPC. 2018. Farm Size and Productivity - Lessons from Recent Literature. https://ispc.cgiar.org/blog/farm-size-and-productivity-lessons-recent-literature

IMMATURE TECHNOLOGIES

While smallholders must compete in markets just the same as other farmers, suitable technologies, applications, software systems and platforms to practice data-driven agriculture and precision farming affordably in small farms are just emerging. Smallholders will continue to operate at a disadvantage until the technologies become more smallholder-ready and the farmers more data-savvy.

The current trend of increasing precision and reducing costs of new farming technologies – especially geospatial and ICTs – will enable small farmers to also practice precision farming and is very positive, but only if other support structures and systems also support them in this shift.

CAPACITY CHALLENGES

Data-driven agriculture is a serious challenge for smallholder farmers due to the quantum of data they will need to process. Given the mix of crops and products on a typical farm, multiplying the data and information needed for each crop or product, per season, reveals the magnitude of data management needed.

The capacities of smallholders to generate and process data and effectively use information are weak. The abilities of current extension systems to step in and take on these data-driven challenges and opportunities are also weak. The lack of capacities in this area in extension systems and other support mechanisms like farmer organizations is important because they have key roles to facilitate data management at an aggregated level. Very few countries and very few farmer organizations have formalized processes to manage all activities related to farm data. As a result, much of the data generated on farms is lost or never captured.

OWNERSHIP CHALLENGES

Besides these capacity challenges, there are wider challenges around ownership. It is simple to believe that whoever farms a piece of land also owns the data and information generated from it. Where land ownership and land cultivation are not in the same hands, this is likely not to be the case. In many places, land is leased or sharecropped for a season or longer by smallholder farmers, without them owning the land. Longer-term data on the farm should be maintained by the owner of the land while current data is maintained by the current user, who changes periodically. In India, for example, a system of soil health cards, which aims to provide farmers the status of soil nutrition on their land and advise on fertilizer interventions for specific crops, is provided to the landowner and not the cultivator, thus raising issues of who owns, and can use and make decisions from the data.

This issue of data ownership for farmers mainly concerns data generated on-farm. It can also be a concern for 'imported' data generated elsewhere for use on-farm where the intellectual property of data from third parties can limit what a farmer does with it.

Relevant legal rights that facilitate access to and use of data at the international, national and subnational level include copyrights, database rights, technical protection measures, trade secrets, and patents and plant breeders' rights, privacy and even tangible property rights.[8] Most of these rights accrue to the intermediaries that invest in data, not the farmers who provide or use data.

EXCLUSION CHALLENGES

As public and private institutions themselves become more data-driven, there is a danger that smallholders, unless included in the data environment, will fall out of the normal systems. Lacking data on their situations, smallholders are unlikely to convince banks, government and credit providers, for example, that they are good investments.

Just as there is a drive in many countries for poor people to join formal financial systems, it is important that smallholders have enough data to thrive and avoid undue risk assessment in the new data-driven systems that are emerging.

One approach might be to create networks of farmers that collect farmers' data and use it to create credit models which allow banks to vet farmers for loans, grants or other public entitlements.

DATA FLOWS AND CONTROL CHALLENGES

In globalized, highly competitive markets, knowing and forecasting agricultural market behavior as early as possible is commercially very advantageous. Access to, and control of, data flows in agri-food systems will be a new battleground. Smallholders, or their representatives, need to make their voices heard.

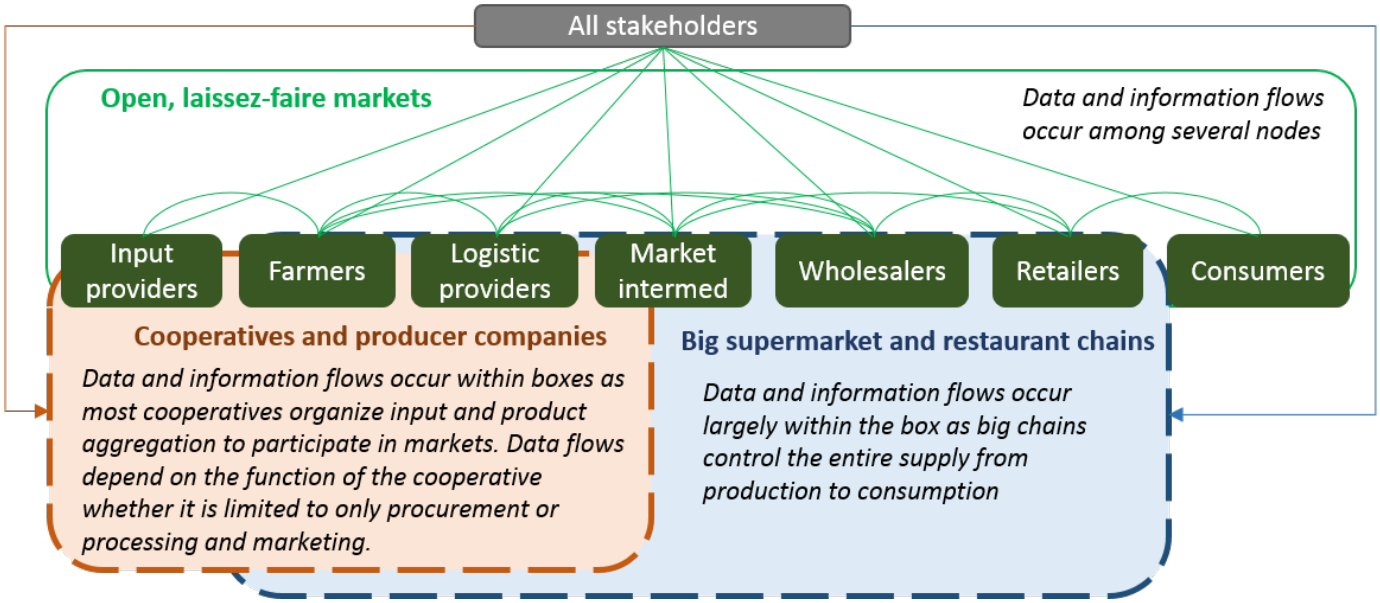
[8]De Beer, J. 2016. Ownership of open data: Governance options for agriculture and nutrition. Wallingford: GODAN. https://f1000research.com/documents/6-1002

Unless something unexpected happens, the trend in economically-developed markets but rapidly spreading in developing countries is for corporate-dominated agri-food chains. These tend to have few value chain nodes that closely link producers, supermarket and consumers. This is unlike the much longer and diverse value chains in developing countries that may include, for instance, producers, collectors, aggregators, wholesalers, public markets, processors, product wholesalers, retailers and consumers.

From a data and information management perspective, having fewer nodes reduces the complexity and makes it easier to improve the efficiencies of information flows. It also makes it easier to manage and control, and potentially exclude, actors in the agri-food system. Thus, agri-food systems such as supermarket chains can be less open with their data and information outside their own systems. Within their systems, the data sharing is controlled on a need to know basis. This data is used to monitor production, resource use and costs, research and innovation and is especially aimed to optimize the agri-food chain for maximum profitability. In such systems, the producers have little decision-making about their farms; they simply produce what is required by the market.

For smallholder farmers to benefit and overcome these challenges, data-driven agriculture must aim at symmetry, balance and equity in the data flow and its use among all actors and stakeholders. This can be achieved through supporting institutions and structures such as policies, legislation, regulations and regulatory mechanisms as well as communications related infrastructure and extension systems that curb asymmetries and enable effective sharing, exchange and use of data by the weaker nodes in the system.

Figure 3 Nodes and Aggregation in Different Agri-Food Systems



Facilitating Smallholder Engagement in Data-Driven Agriculture

Data-driven agriculture thus offers opportunities and poses risks to smallholder farmers. Agri-food systems and farming are becoming more and more knowledge-intensive, but the necessary institutions and support systems, especially extension, advisory and capacity development services, are either not there yet or are not ready to facilitate these changes for the benefit of smallholder farmers. These changes require new capacities among farmers and the institutions that work with them.

Actions should ensure that smallholders are included, given voice and capacities – if necessary through joint efforts, and are enabled to benefit from the undoubted opportunities that enhanced data brings to the sector. There is an urgent need for data-driven applications and systems that are smallholder-ready. Issues and challenges around ownership, control and exclusion must not be forgotten in the drive to roll out the latest technologies and apps.

Groups representing smallholders and producers more generally – associations, federations, cooperatives, social enterprises – must take up these issues and build their own, and their members' capacities. They also need to explore if and how they can become key value adding intermediaries in the various agri-food systems, acting as smart data aggregators and analysts with data importers and exporters.

Making data-driven agriculture smallholder-friendly needs to be guided by critical factors that drive agri-food and data systems. Getting these 'right' will help ensure that data-driven agriculture really benefits smallholder farmers.

Agri-food systems transformation drivers

The speed and effectiveness at which an agricultural system transforms to become more data-driven is a function of several factors. In general, it is critical that the institutional environment allows and encourages data and information to be managed, used, shared and exchanged effectively, equitably and fairly. This environment spans governments, associations of farmers and financial institutions, policies, regulatory frameworks and mechanisms to information and communications related infrastructure and extension systems. These always struggle to keep up in fast-changing digital environments and are often not driven by smallholder concerns. Elements for attention include:

• **Appropriate policies and institutions.**
Protecting the privacy, security, safety and property rights/ownership of all resources, information and finance that flows in the agricultural system is a critical role for policy as they help create a trusted environment that enables many other conditions to happen.

• **Positive incentives**.
Change in this area is rapid and incentives to safeguard and prioritize smallholders need to be devised and taken up. Just as private telecommunications operators are required to provide universal and rural access to less affluent consumers, data service suppliers and operators can extend opportunities and help mitigate risks by adopting ethical guidelines, reinforcing capacities, offering smallholder-friendly services, returning benefits to data owners and being rewarded for this. Innovation is needed in this area. Forms of social certification can also be used as an incentive. Farmer organizations can have an important role here in devising and negotiating such incentives with other actors.

• **Capacities at different levels.**
All actors and stakeholders, especially farmers and small and medium entrepreneurs, need stronger capacities to effectively use data and information. Extended institutional capacities of supporting and advisory organizations are also necessary to manage support systems for data and information sharing and exchange, availability and access to content in the form of relevant data and information,

• **Available, affordable hardware, software and data.**
All the necessary tools including hardware such as sensor equipped farm machinery, software such as applications for the local farming system and cropping pattern, and data needed by the applications needs to be affordable for the smallholder sector. Public policies on one side and negotiations of either governments or farmer organizations with the suppliers have key roles here.

• **Needs-driven data and software**.
More applicable software, models and data can overcome the challenges related to the relevance of data and the usefulness of software and web applications. This also entails more realistic and focused research, not only related to agricultural data guiding the farmer but also presenting

data and information on more user-friendly platforms. It is important that farmers or their representatives are involved in the design of apps, tools, and software. tools, and software.

• **Accessible infrastructure.** Digital services require accessible and affordable connectivity – not only of the Internet and cell phones but also of roads and other transport.

Data system drivers

The data value chain and what it can provide is also fast developing and several dimensions are key to take advantage of. These include 'apps' for farmers that enable localized and specific solutions; open agricultural data and standards that facilitate transparent, equitable and wide use and re-use of data; use of ICTs and data to create jobs and make agriculture attractive to young people; new data handling technical developments like big data, blockchain and Internet of Things that will make data value chains more powerful and transparent; stronger institutions that enable equitable governance of data, locally to globally; and the need for joint actions on data by smallholders and their representatives through farmer organizations, cooperatives, associations, enterprises, etc.

APPS FOR FARMERS

It is a mistake to conclude that developing and providing software such as mobile phone 'apps' is the most critical pathway to introduce digital and data-driven agriculture to farmers. There is an underlying assumption among these application developers that farmers' data and information needs are very simple and that farmers have time and resources to access multiple apps and websites to solve their problems, farm efficiently and participate in markets.

Successful apps must provide localized and specific solutions and therefore use more localized and specific data. Data-driven mobile applications should be built on user requirements and during application development the user should always be placed at the center of design.

At every stage of the farm management cycle, farmers make decisions based on their understanding of their immediate situations, so each step needs an enabling date or ICT solution to systematically accelerate the growth of productivity. These solutions/applications may be specific at one level or on multiple levels, but all should be integrated, contributing to one end.

The most appropriate approach is to develop platforms where all apps can be standardized to be interoperable with sharing of data and information and offer farmers solutions with a basket of options that are location specific. Such platforms will also enable a level playing field for entrepreneurs to offer data and information services.

Development of these tools and platforms must ensure symmetry, balance, equity and fairness in agricultural

OPPORTUNITIES FOR YOUTH

The agriculture sector generally and smallholder farming in particular are not generally attractive propositions for young people in developing countries. However, there are many who argue that ICTs combined with agriculture will help change these perceptions and convince young people to have careers in agriculture.

ICTs, it is suggested, will help create many good job opportunities on and of the farm[9], create opportunities for a new type of e-agriculture entrepreneurship[10]

The apps4Ag database (apps4Ag.org) showcases applications that have been tested and verified. It boosts information dissemination, knowledge exchange, extension and advisory service delivery, farmer engagement, and market access for both agricultural inputs and outputs. It will also develop a framework for assessing the usability and functionality of these applications for agricultural value chain development, mapping the applications along an agricultural value chain framework to facilitate use by value chain actors.

[9] Christiaensen, L. 2017. Can agriculture create job opportunities for youth? World Bank Jobs and Development blog. https://blogs.worldbank.org/jobs/can-agriculture-create-job-opportunities-youth

[10] Technical Centre for Agricultural and Rural Cooperation. 2016. Youth e-agriculture entrepreneurship. ICT Update (83) http://hdl.handle.net/10568/89782

and more take much drudgery out of farming through automation and mechanization. It is not all positive though – increasing use of ICTs on farms will indeed remove some drudgery but is very likely to reduce the need for manual labour on farms, as it did in more developed countries.

There are many examples of initiatives and projects seeking to help young people grasp the opportunities offered by digital agriculture.[11] These include 'hackathons' and support to incubators, accelerators and ICT innovation hubs aiming to boost the number of young people innovating and running sustainable agri-businesses and becoming entrepreneurs.[12]

OPEN DATA

The open agricultural data movement encompasses many forms of agri-food systems, from fully commercial to smallholder subsistence. As these systems evolve, they need primary data from the fundamental unit, the farm, to be open. Both agri-food chain types can benefit from open sharing and exchange of farm data. While the corporate-based systems aim to use the openly shared data to ultimately control and influence markets, the counter movement believes that open sharing of agricultural and related data and information is a beneficial public good.

Two aspects of open data are particularly relevant for farmers. The first is about the data itself and the actual relevance, usability and potential impact of open data.

A recent study[13] on the potential of open data applications for smallholder farmers suggests that, despite the high potential value of open data to smallholder farmers, there are few readily available examples of direct impact on food and nutrition. At least 77 overlapping datasets were identified as applying to or used by smallholders, including government, agricultural law and regulations, government finance data, rural development project data, land use and productivity data, elevation data, value chain data, infrastructure data, meteorological, hydrological, soil, pest management, market and price data. Most current applications with impact however, are using open meteorological and satellite data. Access to other forms of open data is still limited and there are issues with the reliability of data at smallholder resolution. International data sources such as the World Bank and FAO have limited impact in a local smallholder context.

The authors suggest that more scientific and government open data targeting the needs of smallholders is required. It would also be useful to encourage greater collection and sharing of open evaluation data during project implementation.

The second aspect is about the open data approach and ways it can benefit farmers.

Open data approaches, if applied flatly to solve both accessibility and sharing problems and in all types of agri-food systems, are not enough to guarantee equity without a suitable governance framework.

Open access (to data) without capacity to use it effectively only benefits those who have the means to use the data. So, data handling capacities are essential for open data to be fairly and widely exploited.

Further, in data value chains where only some data is open, the ones providing the open data may be exploited. Appropriate systems of checks and balances are necessary to ensure that open data providers are not disadvantaged.

To mitigate these risks, it can also be argued that a fully open data approach could level the playing field for everybody and reduce the risks of monopoly. Alternatively, it may be useful to consider the FAIR framework[14] in which the focus is on clear access rights and appropriate licensing rather than absolute openness and there is special attention to provenance and attribution, aimed at building trust. In the version adopted by the European Commission, this approach is summarized to be "as open as possible, as closed as necessary."[15] Nevertheless, it is important to recognize that not all data will be open,

[11] As an example: Rahman, R. and Fong, J. 2016. Innovate for agriculture: Young ICT entrepreneurs overcoming challenges and transforming agriculture. CTA Success Stories. Wageningen: CTA. http://hdl.handle.net/10568/91708

[12] Technical Centre for Agricultural and Rural Cooperation. 2017. An ICT agripreneurship guide: A path to success for young ACP entrepreneurs. CTA Handbook. Wageningen: CTA http://hdl.handle.net/10568/90136

[13] Jellema A., Meijninger W. and Addison C. 2015. Open data and smallholder food and nutritional security. CTA Working Paper 15/01. Wageningen: CTA.

[14] Wilkinson, M.D. et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data 3: 160018. https://doi.org/10.1038/sdata.2016.18

[15] European Commission Directorate-General for Research and Innovation. 2016. Open research data in Horizon 2020. Infographic. Brussels: European Commission. http://ec.europa.eu/research/press/2016/pdf/opendata-infographic_072016.pdf

to understand the reasons for this, and to devise data sharing and management systems accordingly.

In such open approaches, data should be more open and more accessible; smallholders should also get fair returns from their data. This is not just about farmers selling their data. It could also mean negotiating cheaper and better services. Putting a price on data is not always the best solution as the rich will always have an advantage over the poor, so risks of monopoly can increase.[16] The open data movement can also add value through its work on data standards.

Insufficient data integration is a key barrier in data-driven agriculture, as more and more farm activities rely upon the capture and analysis of multiple sets of data from multiple devices. Standards provide a foundation for this integration and improved abilities to share across platforms will lead to greater efficiencies for all involved. There are clear incentives for farmers and data-driven hardware and software providers to advocate for standardized solutions.

Care needs to be taken to avoid creating multiple standards, often with the mistaken idea that the one being developed is going to be 'the' one standard for all to use. A more promising approach is to work with neutral data standard development organizations, such as the International Standards Organization or AgGateway. They should research and determine if an existing standard already meets any new requirements and whether it can be adopted. They can also propose modified or completely new standards. Implementing or modifying current standards is typically a faster and more successful path than starting from scratch.

BIG DATA, INTERNET OF THINGS AND BLOCKCHAINS

This is a very dynamic and fast-moving area and several new developments around data look set to be at the center of data management and sharing for the coming years. Currently just starting to be used in agriculture, they offer enormous infrastructural opportunities likely to effect whole data chains.

Big data platforms are built to manage and quickly access distributed large-scale data that poses challenges in terms of its volume, velocity, variety and veracity. An important aspect of big data is that it allows precise 'analytics' of this data to discover trends and meaning that can be used to model, build and deliver targeted services. Big Data platforms are essential to handle the amount of data generated by the Internet of Things (IoT), which is all the data coming from all the interconnected 'things' that send data over the Internet.

In agriculture, big data and IoT are currently mainly associated with information collected by sensors, satellites or drones combined with genomic information or climate data, which can all help farmers optimize their farm operations.

Blockchain technology was developed alongside the bitcoin infrastructure as a way to record transactions in secure, transparent and permanent ways. It is particularly important as a mechanism to help build trust along chains of organizations and individuals. "Blockchain is an open, distributed ledger that can record transactions between two parties efficiently and in a verifiable and permanent way."[17] The advantage of blockchain is that there is no one owner of the ledger, transactions cannot be modified by anybody and they are replicated on several servers where the number of identical versions of a transaction on different servers is the means of verification.

In agriculture, blockchain based data management can be very useful especially around data that relates to land and resource use records, purchase and use of pesticides and other harmful agents, traceability and even flows of finance across the whole value chain.

Thus, as an example, payments to different people involved in a final product can be agreed and split using blockchain as the means to verify payments, including to farmers (see box).

This technology provides opportunities to much better understand who benefits in value chains, to track goods and ownership[18], to ensure that income is distributed equitably, and even to create new income-generating possibilities. It is driven by transparency and trust in transactions, as well as the accuracy of data added and shared by all the actors in the chain.

Blockchain is arguably not just a new technology, and not even just a potentially disruptive technology, it is

[16]Tennison, J. 2018. Doesn't open data make data monopolies more powerful? Jeni's Musings blog. https://www.jenitennison.com/2018/01/14/data-monopolies.html

[17]Iansiti, M. and Lakhani, K.R. 2017. The Truth About Blockchain. Harvard Business Review (Jan-Feb): 118-127. https://hbr.org/2017/01/the-truth-about-blockchain

[18]BBC World Service. 2018. Tracing every ingredient on your plate. http://www.bbc.co.uk/programmes/p05z07d7

a foundational technology: "It has the potential to create new foundations for our economic and social systems. But while the impact will be enormous, it will take decades for blockchain to seep into our economic and social infrastructure." Further, "it has the potential to become the system of record for all transactions. If that happens, the economy will once again undergo a radical shift, as new, blockchain-based sources of influence and control emerge."[19]

So far, most blockchain efforts in agriculture are small-scale and starting up. The technology is complex, implementation requires buy-in from many actors, and such potentially transformative changes will take time and good governance to be trusted and taken up widely.

The trust data centers proposed in section 4 seem to be a good institutional placement for data chain actors to come together to establish mechanisms and incentives necessary to govern and coordinate the development of these new technologies in different agri-food systems and value chains.

Blockchain is being used, for example, to connect end consumers, of coffee in Amsterdam with producers in Ethiopia so the consumers can see exactly who produced the coffee, how much they earned from it, and the contributions of other intermediaries in the chain and what they earned. The consumers can also tip the producer as well as the restaurant. In the USA, 'blockchain tomatoes' are being produced with every agronomic variable tracked, reducing spoilage and documenting the supply chain.

Sources:
BBC World Service. 2018. What if you could tip the farmer that grew your coffee? http://www.bbc.co.uk/programmes/p05yxb57
Massa, A. 2017. Someone figured out how to put tomatoes on a blockchain. Bloomberg Technology.

ENABLING INSTITUTIONS FOR DATA-DRIVEN AGRICULTURE

Farmers are only one node, albeit an important one, in agri-food systems, producing and consuming data which, when aggregated across farms, crops, seasons etc., and analyzed such as through big data analytics gains greater value. But, this can also become increasingly hazardous to the interests of farmers' where this symmetry or ability to use information is weak.

Overcoming asymmetries depends on the political and social power of participants in agri-food systems. These are slow to emerge even in developed countries but in developing countries they are almost missing. Those that currently exist or have been developed recently are partial towards satisfying the interests of developed countries and large farmers. The smallholder farmer, the first mile market intermediaries and service providers are generally weak and neglected in these systems. There is an urgent need to relook, renew and develop new institutions and supporting structures for farmers, and more so for smallholder farmers in developing countries for them to shift into practice of data-driven agriculture.

The need for data-driven agriculture has its roots in the need to participate in globally competitive market-oriented agriculture. Development of structures such as policies, strategies, rules, regulations and standards related to agricultural data and information sharing, exchange and flows needs to start at an international level through arrangements, agreements and treaties with structures and mechanisms to implement and regulate them.[20]

Currently, there are no such international efforts. This has been attempted in the sharing and exchange of plant genetic resources based on the recognition of the need to facilitate access to plant genetic resources under a multilateral system, due to the specific features of those genetic resources and their importance for food security. Agricultural and nutrition related data should be treated similarly, contributing to global food security, sustainable development goals and elimination of global hunger and extreme poverty.

Besides plant genetic resources, there are a few other examples of partnerships between stakeholders to improve the governance of certain types of data flows, e.g. through the establishment of national databases.

[19] Iansiti and Lakhani, 2017: op cit.

[20] The Draft UN Declaration of the rights of peasants and other people working in rural areas proposes that [1]) States shall ensure meaningful participation, directly and or through their representative organizations, of peasants and other people working in rural areas in decision-making processes with regard to data identification, research, analysis and interpretation of findings; [2]) Peasants and other people working in rural areas have the right to seek, receive, develop and impart information; and [3]) States shall ensure that relevant information may be adequately disseminated and appropriated by peasants and other people working in rural areas

The closer such initiatives become to international treaties with legal and licensing frameworks, the better they will be able to make data flows more equitable.

The development of country-level policies, strategies, rules, regulations and use of standards is also important to support and supplement international initiatives.

EMERGING ROLES FOR JOINT ACTION

Individual actions by smallholders will have little impact in this data-driven environment. Joint action by smallholders offers ways to jointly safeguard their own data, maximize returns in value chains, and best exploit the potential of third-party services and data offerings. One of the most promising routes to this is working through farmer-representing organizations that promote their members' interests.

Data-driven services and products are thus coming to be seen as promising mechanisms that such farmer organizations – cooperatives, associations, enterprises, etc., - can use to better serve the interests of their members.

Recent work[21] by the Technical Centre for Agricultural and Rural Cooperation identified a number of enabling factors and tools necessary for these organizations to help achieve the desired impact for smallholders, including: international, national and sectoral regulatory frameworks; adequate ICT infrastructure, capacities and access; effective data policies and ethics – at national and institutional level; access to open or FAIR data; sufficient capacities to understand, use and share data – for individuals and organizations; communication channels among the various actors; trust –necessary for farmers to rely on the services; as well as various tools to collect, monitor, track, measure and deliver data.

Systems to generate and manage farmer registration and profiles are seen to be especially critical to deliver relevant and timely services and products where they are most needed. Experiences suggest that such farmer profiling allows the farmer organizations to connect better with their members and deal with third parties. For this to work, it is critical that farmer organizations understand (and safeguard) the value of the data they have. The farmers also need to understand the potential benefit of their data, in that it gives them a stronger voice and power.

[21] See CTA. 2017. Data-driven products and services for farmers organizations.

[22] Boyera, S., Addison, C. and Msengezi, C. 2017. Farmer profiling: Making data work for smallholder farmers. CTA Working Paper 17/09. Wageningen: CTA


time
10:24
dcf
date
20. 6.
forecast
in
19.3 °C
50 % rh
out
16.8 °C
1
SET
IN
OUT
CH

Developing Data Ecosystems for Smallholder Farmers

This section draws on findings from a four-day training session and symposium on Farmers' Access to Data convened by the Global Forum on Agricultural Research and Innovation as part of a set of actions on farmers' rights to data delivered with GODAN and CTA.[23]

Course participants, mostly farmers and representatives of farmers' organizations, developed a vision to enable smallholder farmers to become successful business owners by using agricultural data. They used the term 'agripreneur' to help change the mindset of managing a plot to managing a business. Developing a data ecosystem to support farmers requires:

1. **Aggregating farmer data through joint action that empowers and gives voice to farmers**

Cooperatives, producer organizations, farmer organizations and similar farmer-representing organizations have important roles to promote and enable aggregation of farmer initiatives, including managing, sharing and using data and information.

Such shared actions by smallholders offer ways to jointly exploit and safeguard their own data, maximize returns in markets, and best exploit the potential of third-party services and data offerings. From the base of a group or cooperative, farmers can push leaders to implement data rights and policies and they can provide a better service to consumers, restaurants, markets and other retailers.

Finding, testing and implementing sustainable business models for farmers and farmer organizations to benefit from data is an urgent action. Farmer-representing organizations need to transition towards digitally-smart organizations or 'data cooperatives' that are able to broker data-driven interactions between and among their members and with external organizations of all types.

Besides any local aggregation of farmers, the 'virtual' aggregation of farms could facilitate more synchronized access to farm inputs, processes, outputs and logistics to participate in markets through use of ICTs.

2. **Establishing trust centers, mechanisms and platforms to enable and regulate open data sharing at different levels**

Setting up trust centers and platforms at different levels (commodity-specific, value-chain segment-specific, regional, national…) would help bring together and facilitate necessary joint efforts and actions to improve the ways that data is accessed, shared and applied in agriculture and nutrition. It would also help deliver international agreements with incentive, regulatory and enforcement mechanisms to share data at various levels and among different actors across the data value chain, from plot, farm, farming system, region, national to global agricultural and related systems.

Trust centers could be created around farmer-representing organizations or consortia of farmers and industry representatives along a segment of the value chain, perhaps facilitated and supported by the public sector. They could be ideal locations for discussion among public and private sector actors, among farmers and big industry, devising incentives to achieve mutual benefits.

Trust centers could take on several important roles and issues, such as:

- **Governance and incentives**
Facilitate inclusive governance of flows of data, information, knowledge, skills and technology; Promote appropriate legal and licensing frameworks (as an example, a 'data commons' has been proposed with a licensing model in which farmers benefit from the datasets they contribute to[24]); and advocate for effective incentive and other schemes to foster trust, control and transparency. This could include using social certification of data services to promote best practices or setting up blockchain networks to reinforce security and trust in value chains.

- **Integration across the value chain**
Test new business models and tools to better integrate governments, farmers, banks, insurance companies, market intermediaries, cooperatives etc. for data-driven participation in agri-food systems; and facilitate platforms for interaction among different actors in the agri-food chain. Trust centers could be also a good place to experiment with pre-competitive spaces for the private sector to share data and technologies.

- **Data ecosystem support**
Develop new forms of advisory and support systems to build needed capacities to use data-driven information, knowledge, skills and technology; Develop inclusive standards for data collection, sharing, interoperability, accessibility and accuracy; promote open technologies for

[23] Pesce V, and Berne, D. 2017. Learning about data for farmers and how it can help to cross the donga. https://article.wn.com/view/2017/12/20/Learning_about_data_for_farmers_and_how_it_can_help_to_cross

[24] Baarbé, J., Blom, M. and Beer, J. de. 2017. A data commons for food security. Working Paper 7. Open AIR. http://www.openair.org.za/publications/a-data-commons-for-food-security

farming and on-farm processing of farm products as also for data and information; and devise and perhaps govern data management and sharing platforms that ensure transparency, recognition of ownership and traceability of data.

• **Innovation and experimentation**: Test and disseminate innovative solutions to ensure efficient and fair sharing of data such as: state-of-the-art technologies to ensure efficiency (like big data platforms) mechanisms to ensure transparency and recognition of ownership (like blockchain); engage technology providers and incubators to become part of the trust center or negotiate outsourced technology solutions.

Establishing such trust centers requires investments and commitments from different local actors. Well thought out proposals and business cases from groups of actors well positioned to play these roles, illustrating governance options to avoid duplication and strategies for sustainability, together with a clear case will be needed to attract investors.

3. **Developing international agreements to govern data access, ownership and flows**

These are needed if agricultural and nutrition related data is to be treated as a universal public good related to global food security, sustainable development goals and elimination of global hunger and extreme poverty. Agreements that oblige countries and institutions to follow certain behavior would demonstrate the benefits and opportunities from sharing data and facilitate equitable, fair, ethical flows and uses of agricultural and nutritional data.

Annex 1. Data in crop farming cycles

The data and information that farmers use is typically event driven in a farming or seasonal cycle, with cycles varying across commodities and farming systems. Phases of such cycles can be identified with their differing data needs for pre-planting planning, acquiring and testing inputs (including any loans or credit), planting (for an annual crop), cultivation, harvest, post-harvest processing, marketing, etc.

To follow the crops example, for the pre-planting planning phase, farmers may need to know the archived cropping pattern of the land, weather patterns and climate, soil potential and conditions, rainfall forecasts and irrigation availability in the cropping season, forecasted prices of inputs, time to harvest the crop, possible yield, forecast of market price of the produce and possible return on investment to decide the crops in a particular season for the farm.

The planting-related data includes data on land preparation according to the requirements of the chosen crop, use and availability of farm machinery and farm labor, seed quantity needed, fertilizer and manure needs, climatic data especially rainfall and temperatures as well as irrigation schedules, etc.

At cultivation phase, the farmer may need data to monitor the crop such as for growth, soil humidity, pest and weed density, use of pesticides and herbicides etc.

At the harvest and post-harvest phase, the farmer needs data and information on the right stage of the crop to harvest, availability of farm machinery and labor, storing and packing the main product and the byproducts, market price of the products, transport availability and cost etc.

For the data to be relevant and useful its context, it needs to be associated. Data loses its context if separated from the information a farmer needs and uses. This is sometimes seen, for example, when farmers are given raw weather data without the context of what the data means for a farming operation. For example, farmers do not see usefulness in just rainfall data. Its use must be interpreted in the agronomic context of whether the pattern of rainfall is adequate to initiate sowing of a particular crop at a particular time. The context of rainfall data may be different if the crop, for example is rice or cotton. Data should therefore become information in an adaptive context.

Even within a crop, the variety, cost and availability of inputs etc. play a very critical role in interpreting data such as of rainfall for an event in the farming cycle. For farmers, the use of data, information and knowledge are and will be intertwined and interrelated and cannot be treated individually and independently.

Table 1 Illustrative data and information required by a farmer for crop farming (with some local parameters specific to Gujarat, India)

Decision / Data Step

Typical Source of Data: Localized

Typical Source of Data: Imported

Purpose of cultivating a particular crop (Profit, Market participation, Risk management, Soil improvement, Subsistence)	How to use, Price and Availability with Local Source of Fertilisers, Pesticides, Weedicides, Farm Machinery, Fuel
Desired characteristics of Crop and Variety sown such as characteristic of product, resistance to drought, disease, pests etc.	PHI, NHI, Safety and Withdrawal period etc. of pesticide, weedicide etc before harvest
Farm, Field and plot layout and size	Weather Forecasts (Agro-meteorological such as rainfall and its pattern, Temperature, Days cloudy, chance of frost etc.)
Subsidies, Discounts and Support price for the crop	Irrigation Type
Average time to prepare 1 hectare for sowing	Time in Days of Different Crop Maturity Stages
Seeding Rate per Hectare	Recommended Machinery
Optimum Plant Density per Hectare	Approximate Labour required For Pre-Planting Preparation, Planting, Weeding, Spraying etc.)
Recommended Date for Sowing	Farm Machinery needed for inter-cultivation and their availability
Days to harvest the crop from date of sowing	Seeding/Transplantation Method
Seed availability in Locality (Taluk/Village Cluster)	Time needed to Seed/Transplant per Hectare in Man days
Approximate Cost of Seed per Kg	Interculture Date (Repeated)
Potential Yield in Kg for Full crop	Fertilizer application
Recommended Soil Type	Water required and periodicity of irrigation
Recommended Soil Humidity % at Planting	Soil Temperature (Repeated)
Recommended Soil Temperature at Planting	Soil pH (Repeated)
Recommended soil Ph at planting	Soil Humidity (Repeated)
Recommended Soil N (kgs per hectare)	Approximate Stage of Crop when a particular pest / disease / weed attacks the crop
Recommended Soil P (Kgs per Hectare)	Machinery for Harvest
Recommended Soil K (Kgs per Hectare)	Method of Harvest (Single/Multiple Picking as in Castor/Cotton)
Required Micronutrients (Gms/Kg) per hectare	Estimated Labor required for Harvesting in days per Hectare
Required Micronutrients (Gms/Kg) per hectare	Storage for Grain
Recommended Quantity of Farm Yard Manure	Average Yield of the crop per hectare realised by Farmer
Recommended Vermicompost	Postharvest Processing
Proposed Pesticide Usage List for the Crop	Postharvest Packaging
How to use, Price and Availability with Local Source of Fertilisers, Pesticides, Weedicides, Farm Machinery, Fuel	

Publishers

Global Forum on Agricultural Research and Innovation (GFAR)
c/o United Nations Food and Agricultural Organization
Viale delle Terme di Caracalla Roma 00153, Italy
E-mail: gfar-secretariat@fao.org
Tel: +39 06 5705 3413
www.gfar.net

Global Open Data for Agriculture and Nutrition (GODAN)
c/o CABI
Nosworthy Way, Wallingford, Oxfordshire OX10 8DE, United Kingdom
Email: enquiries@godan.info
Tel: +44 (0) 1491 832111
www.godan.info

Technical Centre for Agricultural and Rural Cooperation (CTA)
P.O. Box 380, 6700 AJ Wageningen, The Netherlands
E-mail: cta@cta.int
Tel: +31 (0) 317 467 100
www.cta.int

Acknowledgments

Special thanks to the following organizations for supporting the work of some of the authors:
Earth Observation Science & Information Technology of CSIR South Africa (www.csir.co.za)
Land Use and Soil Management Directorate of the Department of Agriculture Forestry and Fisheries of South Africa (www.daff.gov.za)
Global Institute for Food Security (www.gifs.ca)

Disclaimer

This document has been produced with the financial assistance of the European Union. The contents of this document are the sole responsibility of the authors and the publishers and can under no circumstances be regarded as reflecting the position of the European Union.

License